



Version with Markings to
Show Changes Made

DESCRIPTION

PROBE FOR PROTEIN-PROTEIN INTERACTION ANALYSIS, AND METHOD OF
~~USING IT FOR ANALYSIS OF PROTEIN-PROTEIN INTERACTIONS~~ USING THE
SAME

This application is a 371 of PCT/JP00/009348, filed December 27, 2000.

TECHNICAL FIELD

The invention of this application relates to a probe for protein-protein interaction analysis, and to a method of using such probe for the analysis of protein-protein interactions using such a probe. More precisely specifically, the invention of this application relates to a probe for protein-protein interaction analysis and to a method of using it for the analysis of protein-protein interactions using such a probe, which enables enable accurate and simple analysis of protein-protein interactions in all living cells.

BACKGROUND ART

~~It is well known that protein-protein~~ Protein-protein interactions are known to play key roles in structural and functional organization an important role with respect to structure and functions of living cells.

Many unsolved problems currently studied in molecular biology and biochemistry, such as for gene transcription mechanism and intracellular information transmission signaling, are related to protein-protein interactions.

Some of the problems in the art field of molecular biology and biochemistry have been gradually heretofore solved by the development of protein library screening techniques such as a two-hybrid strategy, in which a library of "target" proteins is screened for interactions with "bait" proteins method (Chien, C.

T., Bartel, P. L., Sternglanz, R., Fields, S., *Proc. Natl. Acad. Sci. USA* 1991, 88, 9578-9582; Fields, S., Song, O., *Nature* 1989, 340, 245-246) by which screening is conducted based on interaction of a "bait" protein and a "prey" protein in a protein library. The two-hybrid method has been suggested as an effective method that facilitates the identification of potential protein-protein interactions and has been proposed as an effective method for the generation of protein interaction maps candidate molecules for protein-protein interaction and also creates new protein-protein interaction maps (Flores, A., Briand, J. F., Gadal, O., Andrau, J. C., Rubbi, L., Mullem, V., Boschiero, C., Goussot, M., Marck, C., Carles, C., Thuriaus, P., Sentenac, A., Werner, M., *Proc. Natl. Acad. Sci. USA* 1999, 96, 7815-7820; Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M., Yamamoto, K., Kuhara, S., Sakaki, Y., *Proc. Natl. Acad. Sci. USA* 1999, 97, 1143-1147; Walhout, A. J. M., Sordella, R., Lu, X., Hartley, J. L., Temple, G. F., Brasch, M. A., Thierry-Mieg, N., Vidal, M., *Science* 2000, 287, 116-122). However, the problem associated with the two-hybrid method has a problem in is that it is applicable only to the detectable analyzable protein interactions that occur around close to the reporter gene in cell nuclei but not to any other ordinary protein-protein interactions and, therefore, this method lacks generality.

Reliability is yet another problem associated with the two-hybrid method is that its reliability is low. Therefore, every assay according to the method must be followed by an additional confirmation test using proteins of model cells or animals of which the functional data are known, and a confirmation test using proteins with known functions has to be conducted by employing model cells or animals for every assay (Walhout, A. J. M., et al., *Science* 1999, 287, 116-122).

Given that situation Accordingly, novel methods for other

protein-protein interaction assay strategies have been developed. One is analysis have been suggested, those methods including a ubiquitin split protein sensor (USPS) method by which comprises reconstructing N- and C-terminal ubiquitins are reconstructed through protein-protein interaction followed by activating the reactions with proteins and a nuclear-localized reporter is activated by cleavage of a through transcription factor cleavage (Dunnwald, M., Varshavsky, A., Johnsson, N., *Mol. Biol. Cell* 1999, 10, 329-344; Johnsson, N., Varshavsky, A., *Proc. Natl. Acad. Sci. USA* 1998, 95, 5187-5192; Stagljar, I., Korostensky, C., Johnsson, N., Heesen, S., *Proc. Natl. Acad. Sci. USA* 1998, 95, 5187-5192; and another is an SOS recruit system, in which a catalyst domain is brought near close to an intended a membrane localization domain through protein-protein interaction, thereby reconstructing a guanine exchange factor (GEF) or Ras is reconstructed, which then and it further complements yeast causing having a temperature-sensitive mutation in mutated yeast GEF (Aronheim, A., *Nucleic Acids Res.* 1997, 25, 3373-3374; Aronheim, A., Zandi, E., Hennemann, H., Elledge, S. J., Karin, M., *Mol. Cell. Biol.* 1997, 17, 3094-3102; Broder, Y. C., Katz, S., Aronheim, A., *Curr. Biol.* 1998, 8, 1121-1124).

As a more general approach, a split enzyme technology method has been reported (Rossi, F., Charlton, C. A. and Blau, H. M., *Proc. Natl. Acad. Sci. USA* 1997, 94, 8405-8410; Remy, I., Michnick, S. W., *Proc. Natl. Acad. Sci. USA* 1999, 96, 5394-5399; Pelletier, J. N., Arndt, K. M., Pluckthun, A., Michnick, S. W., *Nature Biotech.* 1999, 17, 683-690). It is said that, according to the With this method, a split enzyme is reconstructed through protein-protein interaction therein, and its enzymatic activity is restored. Further, it is that the The activity of the thus-reconstructed enzyme can be identified measured by its fungi or cellular phenotype or by its a fluorescent enzyme substrate

analogues.

These various methods produce accurate assay data in some degree and are well suited for have a comparatively high accuracy and attain the object of assaying interactions between intracellular proteins and or membrane proximal proteins, however but they still contain problems in that they can be are applicable to only appropriately designed cells. Another such problem is that their sufficient accuracy is still unsatisfactory and their sensitivity is not sufficient cannot be obtained. Still another problem is that such methods require different a variety of substrates to conduct the analysis and are therefore troublesome.

At present Therefore, a general method applicable to any protein that enables or probe suitable for accurate and simple analysis of various protein-protein interactions or an all-purpose probe for such method is not known and applicable to any protein has not yet been developed.

The invention of the present application has been made in consideration of the above described situation With the foregoing in view, it is an object of the invention of the present application to solve the problems of the prior art and to provide a an all purpose probe applicable for various protein-protein interaction analysis and a method of protein-protein interaction analysis using such probe, which enables highly accurate and simple assay analysis of the interaction of proteins protein-protein interactions, and to a method for the analysis of protein-protein interactions using such a probe.

DISCLOSURE SUMMARY OF THE INVENTION

In order to solve the above-mentioned problems, the invention of this application firstly provides a probe for analyzing protein-protein interaction between two proteins analysis, this probe serving to analyze the interaction between

two proteins, wherein protein splicing is induced by protein-protein interaction, thereby regenerating and a physicochemically or biochemically detectable protein is regenerated.

Secondly, the invention provides the probe for protein-protein interaction analysis, consiting of comprising two probes ~~that are~~: a probe a"a" which comprises an N-terminal polypeptide of an intein and an N-terminal polypeptide of a labeled protein, and a probe b"b" which comprises a C-terminal polypeptide of the intein and a C-terminal polypeptide of the labeled protein.

Thirdly, the invention provides the probe for protein-protein interaction analysis, wherein the C-terminal of probe a"a" and the N-terminal of probe b"b" each has a linker sequence.

Fourthly, the present invention provides the probe for protein-protein interaction analysis, wherein the intein is an endonuclease derived from yeast VMA; and fifthly, the invention provides the probe for protein-protein interaction analysis, wherein the intein is DnaE derived from cyanobacterium.

Sixthly, the invention provides the probe for protein-protein interaction analysis, wherein the ~~labeled protein~~ indicator protein (e.g., labeled protein) is a fluorescent protein; and seventhly, the invention provides the probe for protein-protein interaction analysis, wherein the fluorescent protein is a green fluorescent protein (GFP).

Eighthly, the invention provides the probe for protein-protein interaction analysis, wherein the ~~labeled protein~~ indicator protein is an emission-catalyzing enzyme; and ninthly, the invention provides the probe for protein-protein interaction analysis, wherein the emission-catalyzing enzyme is a luciferase.

Tenthly, the invention provides a method for analyzing the analysis of protein-protein interaction interactions, comprising ~~making the steps of causing~~ a protein linked with probe a as described in claims 2 to 9 and to the probe "a" of any of the above-described second to ninth claims to be co-present with a

protein linked with to the probe b "b" as described in claims 2 to 9 coexist in a system; of any of the above-described second to ninth claims, and detecting the a signal emitted by of the labeled protein.

Eleventhly, the invention provides the method of protein-protein interaction analysis, wherein the protein linked to the probe "a" is caused to be co-present with the protein linked to the probe "b" by introducing a polynucleotide that expresses the probe of any one of claims 1 to 9 is introduced into a eucaryotic cell, thereby making the probe a-linked protein and the probe b-linked protein coexist in the cell for protein-protein interaction analysis of any of the above-described first to ninth claims into an eukaryotic cell.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows is a schematic drawing illustrating an example of the structure and functional operation principle of the probe for protein-protein interaction analysis of the present invention. Here, (I) shows co-existing two probes illustrates co-presence of protein pairs and probes; (II) shows illustrates protein-protein interaction; (III) shows illustrates protein splicing; and (IV) shows illustrates linkage of labeled protein indicator protein linkage. Furthermore, (1a) indicates probe a denotes the probe "a" for protein-protein interaction analysis; (1b) indicates probe b denotes the probe "b" for protein-protein interaction analysis; (2a) indicates denotes the N-terminal polypeptide of an intein; (2b) indicates denotes the C-terminal polypeptide of the intein; and (3) indicates denotes a labeled protein. Further In particular, (3a) indicates denotes the N-terminal polypeptide of the labeled protein; and (3b) indicates denotes the C-terminal polypeptide of the labeled protein. (4a) indicates denotes protein (or a protein site) A; and (4b) indicates denotes protein (or protein site) B.

Fig. 2 shows amino acid residues substituted between I124 and I129 with respect to the structure of the N-terminal polypeptide ~~structure~~ of EGFP in the Example, for which the 124th to 129th amino acid residues have been replaced of the present invention.

Fig. 3 shows the data results of SDS-PAGE analysis of proteins expressed in *E. coli* and transformed with pGEX-NVC. Here, lanes (a) to (c) are Coomassie Blue-stained SDS-PAGE; lane (a) indicates the protein molecular mass standard (Novagen) with the molecular mass thereof (kDa); lane (b) indicates the crude ~~extract product before purification~~ a GST-affinity column; and lane (c) indicates the sample purified by the GST-affinity column. Lanes (d) and (e) show Western blotting analysis of a crude lysate ~~using antibodies specific to VDE (d) and to GFP (e);~~ (d) shows the analysis results obtained with the antibodies specific to VDE and (e) shows the results of analysis obtained with the antibodies specific to GFP (e).

Fig. 4 shows is a schematic representation of the structure of a plasmid (probe <1> for protein-protein interaction analysis) constructed in the Example of the present invention. Here, the restriction endonuclease splicing sites are shown above ~~each bar~~ the bars, the dotted lines indicate the endonuclease domain in the VDE intein[[;]], the open bars indicate the intein VDE intein with no endonuclease[[;]], the black bars indicate GST labeling[[;]], and the ~~shadowed bars indicate hatching indicates~~ His labeling. The Furthermore, the lattice patterns indicate a linker[[;]], "stop" indicates a translation termination codon[[;]], and "start" indicates a translation initiation codon.

Fig. 5 shows the data of splicing results induced by obtained with interaction of protein CaM and protein M13 in the Example of the present invention. Here, (a) indicates pET_NVC Δ SD (C/M) (SDS-PAGE); (b) indicates pET_NVC Δ SD (M/C); (c) indicates pET_NVC Δ SD (/) gene 1; (d) indicates pET_NVC Δ SD linker (C/M); (e)

indicated pET_NVC Δ SD linker (M/C); and (f) indicates pET_NVC Δ SD linker (/). (A) shows the results of Western blotting ~~labeled obtained with anti-His labeling[();]~~, and (B) shows the ~~data of results of Western blotting labeled obtained with anti-GFP labeling.~~

Fig. 6 shows the fluorescence spectra of crude products from *E. coli* ~~carrying any one of the having~~ plasmids (probe <1> for protein-protein interaction analysis) constructed in the Example of the present invention ~~Excitation is at (excitation: 470 nm, the excitation band width is: 5.0 nm; and the emission band width is: 5.0 nm)~~. (a) indicates NVC Δ SD linker (/); (b) indicates NVC Δ SD (C/M); and (c) indicates NVC Δ SD linker (C/M).

Fig. 7 shows the structure of a plasmid used in the Example of the present invention. Here, the dotted lines indicate an intracellular ribosome ~~extein~~ entry site (IRES); the cDNAs of pLucAll, pLucN and pLucC are inserted into pcDNA3.1 (+); "stop" "Stop" and "start" "Start" indicate translation termination and initiation codons, respectively.

Fig. 8 shows the emission intensity of LucN alone, LucC alone and LucAll in the Example of this invention.

Fig. 9 shows the ~~time-dependent enhancement of protein splicing increase by insulin~~ in the Example of this invention.

Fig. 10 shows the influence of amino acid mutation ~~on protein splicing~~ in the Example of this invention.

Fig. 11 shows the insulin concentration dependency of ~~Lum-F~~ Lum-F emission intensity in the Example of this invention.

~~BEST MODE FOR CARRYING OUT DETAILED DESCRIPTION OF THE INVENTION~~

The probe for protein-protein interaction analysis of the present invention is based on the principle that the interaction of two proteins, ~~each being linked to the probes a probe, induces splicing thereby regenerating the labeled protein split and attached to the two probes, to emit a signal, the sites of split~~

labeled proteins contained in the probes are linked and regenerate the labeled proteins, and a signal is emitted.

Protein splicing is a process wherein the internal protein segment (intein) is excised from a translated protein. In this process, the excision of intein is accompanied by the ligation of flanking sequences (exteins) (Gimble, F. S., *Sci. Biol.* 1998, 5, R251-256).

The With the probe for protein-protein interaction analysis of the present invention, protein-protein interaction analysis is provided conducted by ingeniously applying the such self-excision of inteins to protein-protein interaction analysis.

In other words, the probe for protein-protein interaction analysis of the present invention ~~consist~~ consists of two probes ~~a~~ a "a" and ~~b~~ b", each ~~separately probe being connected to two different proteins a respective protein.~~ Fig. 1 is a schematic representation of the principle of the probe for protein-protein interaction analysis of the present invention.

Of the two probes ~~a~~ a "a" (1a) and ~~b~~ b" (1b), probe ~~a~~ a "a" (1a) ~~contains comprises~~ the N-terminal polypeptide of an intein (2a) and the N-terminal polypeptide of a ~~labeled protein indicator protein~~ (3a)[[;]], and probe ~~b~~ b" (1b) ~~contains comprises~~ the C-terminal polypeptide of the intein (2b) and the C-terminal polypeptide of the ~~labeled protein indicator protein~~ (3b).

The two probes ~~a~~ a "a" (1a) and ~~b~~ b" (1b) are ~~separately connected linked~~ to the respective proteins A (4a) and B (4b), for which interaction is to be analyzed, to detect analyze the protein A - protein B interaction. Hence, when protein A (4a) and protein B (4b) ~~each connected linked~~ to probe ~~a~~ a "a" (1a) and probe ~~b~~ b" (1b), ~~respectively, coexist are co-present~~ (I) and when the two proteins (4a, 4b) interact with each other (II), the intein (2) is excised through splicing (III). As a result, the ~~labeled protein indicator protein~~ sites (3a, 3b) ~~bonded bonding~~

to the intein (2) are ligated (IV), thereby enabling the confirmation of the protein-protein interaction through detection of the signal resulting from the labeled protein (3). Therefore, if the obtained indicator protein (3) is analyzed, the protein-protein interaction can be confirmed.

In this process, if protein A (4a) does not interact with protein B (4b), ~~splicing excision~~ of intein (III) ~~and splicing of exteins does do~~ not occur; therefore, the ~~labeled protein indicator protein~~ (3) is not regenerated (IV) and cannot be detected.

For the probe for protein-protein interaction analysis of the present invention, probes ~~a~~ a and ~~b~~ b (1a and 1b) may each be composed only of the intein polypeptide and the ~~labeled protein indicator protein~~ polypeptide, but may also additionally contain other parts such as a linker sequence or the like. Accordingly, in probes ~~a~~ a and ~~b~~ b (1a and 1b), the intein polypeptides (2a, 2b) and the ~~labeled protein indicator protein~~ polypeptides (3a, 3b) may be directly bonded bound to each other, or may be bonded bound via ~~any other~~ a polypeptide such as a linker sequence.

~~The intein may be well-known inteins derived from various known organisms can be employed as the intein. It includes, for example, Suitable examples include those derived from eucaryotes eukaryotes~~ such as *Saccharomyces cerevisiae* (yeast) Sce VMA, *Candida tropiannis* (*Candida* fungi) Ctr VMA; those derived from eubacteria such as *Mycobacterium tuberculosis* (tubercle bacillus) Mtu recA; and those derived from paleobacteria such as *Thermoplasma asidophilum* Tac VMA. In addition, *Cyanobacterium synechocystis* sp. (cyanobacterium) DnaE may also serve as the intein ~~for use in the invention~~.

When protein A (4a) interacts with protein B (4b) in the presence of the probes (1a, 1b) for protein-protein interaction analysis ~~(1a, 1b)~~ of the present invention, the inteins

~~facilitate an autocatalytic splicing reaction are preferably site-specific endonucleases to provide for autocatalytic intein excision.~~

Specifically, intein derived from yeast VMA and Ssp DnaE intein derived from cyanobacterium are ~~preferable the preferred examples~~. In yeast VMA, the nascent translation product, 120-kDa VMA1, catalyzes protein splicing to give a 70-kDa H^+ -ATPase H^+ -ATPase sub-unit and a 50-kDa site-specific endonuclease (VDE, or PI-SceI). This VDE is preferably used as the intein sites (2a, 2b) in the probes (1a, 1b) for protein-protein interaction analysis ~~(1a, 1b)~~. For the Ssp DnaE derived from cyanobacterium, the DNA sequence of the strain PCC6803 has been ~~determined clarified~~ (the N-terminal has 123 amino acid residues and the C-terminal has 36 amino acid residues); ~~further moreover, Ssp DnaE intein it is a natural split intein[[],]~~ and ~~is known that the ligation of N- and C-extein occurs is known to occur therein~~ (Wu, H., Hu, Z., Liu, X., XU., M., Q., *Proc. Natl. Acad. Sci. USA* 1998, 95, 9226-9231). Therefore, Ssp DnaE intein is easy to handle~~[[],]~~ and preferred as the intein sites (2a, 2b) in the present probes (1a, 1b) for protein-protein interaction analysis ~~(1a, 1b)~~.

Of the various inteins mentioned above, the cyanobacterium-derived Ssp DnaE intein is ~~most preferable for enabling highly sensitive detection of preferred, because the protein-protein interactions in mammal cells can also be detected with high sensitivity. It is needless~~ Needless to say, that any other known or novel inteins may also be employed in the invention.

For effective excisions of intein and splicing (III) ~~with of the probes for protein-protein interaction analysis (1a, 1b) of the present invention, probes 1a and 1b must for protein-protein interaction analysis have to be correctly folded and sites have to be correctly aligned, so that the two sites participating in protein splicing are be adjacent to each other~~

and are correctly aligned (Duan, X., Gimble, F. S. and Quiocho, F. A., *Cell* 1997, 89, 555-564). Accordingly, as the inteins, those derived from organisms may be directly used or they may be suitably modified for facilitating the intended intein splicing by, designed so as to facilitate the splicing, for example, partly substituting by transforming or deleting some of the amino acid residues of the organisms derived intein with any other residue, partly deleting amino acid residues, or by introducing a suitable linker sequence thereinto.

For example, for the above-mentioned VDE, it is known that its mutant specifically designed by deletion of a mutant obtained by deleting the endonuclease domain and substitution substituting with a foldable flexible dodecapeptide linker shows an increased is known to demonstrate a high splicing activity (Cooper, A. A., Chen, Y. J., Lindorfer, M. A., Stevens, T. H., *EMBO J.* 1993, 12, 2575-2583; Chong, S., Xu, M. -Q., *J. Biol. Chem.* 1997, 272, 15587-15590). On the other hand, the cyanobacterium-derived Ssp DnaE is a natural cleaved naturally split intein, as mentioned above, and is known to lead to fold correctly folded splicing and splice the exteins., Therefore, The above-mentioned linker sequence may or may not be introduced into dnaE it.

As described hereinabove, the The above-described intein is split into the N-terminal polypeptide (2a) and the C-terminal polypeptide (2b). In the probes (1a, 1b) for protein-protein interaction analysis of the present invention, the split polypeptides (2a, 2b) are bonded bound to the N-terminal polypeptide (3a) of a labeled protein indicator protein and the C-terminal polypeptide (3b), respectively, to construct become a probe.

On the other hand, for in the probes (1a, 1b) for protein-protein interaction analysis (1a, 1b) of the present invention, the labeled protein indicator protein sites to be used undergo direct peptide bonding and link to each other are directly

peptide bonded and linked (IV) through the interaction of proteins A and B (II), followed by the intein splicing from probes a and b (1a, 1b) (III) when the intein is spliced out and probes "a" and "b" (1a, 1b) (III) are spliced by the interaction (II) of proteins A and B. The labeled protein indicator protein (3) may be any protein that can be detected after protein analyzed again by the linkage (IV). For example, fluorescent protein and emission-catalyzing enzyme are preferable preferably employed. As a fluorescent protein Fluorescent proteins such as green fluorescent protein (GFP) which emits are preferred because they emit light after protein linkage, enabling visual detection, is preferable and can be analyzed visually. For the emission-catalyzing enzyme Emission-catalyzing enzymes such as luciferase that forms an activity center after protein are also preferred because they form active centers after linkage and emits emit light that is can be easily detected is preferred with a luminometer. For the split N- and C-terminals of the luciferase to give individually no fluorescence when separated and to restore its emission the activity again after its two terminals are recombined bonding, the luciferase must be split so that its activity active center is divided into two. It is known that the enzyme The luciferase enzyme is known to be folded into two domains that sandwich a broad activity center containing region therebetween, on is comprising an active center, one being a large N-terminal domain comprising one β -barrel and two β -sheets and the other is a C-terminal site (Waud, J. P., Sala-Newby, G. B., Matthews, S. B., Campbell, A. K., *Biochim. Biophys. Acta* 1996, 1292, 89-98; Conti, E. Franks, N. P., Brick, P., *Structure* 1996, 4, 287-298). Accordingly, it is desirable preferred that the enzyme luciferase is be split into two 3a and 3b[[],] at the flexible side through which site where the two domains are linked to each other. For With respect to the enzyme luciferase enzyme, it is further also known that its the N-terminal polypeptide

(3a), to which the N-terminal polypeptide of the intein (2a) is to be bonded, has a cysteine (Cys) residue in the bonding site with the N-terminal polypeptide(2a) of the intein, and further has a tyrosine (Tyr) residue ~~at~~ upstream to it of -1 position thereof, and has alanine (Ala) and phenylalanine (Phe) residues in the -3 and -4 upstream to it positions, respectively; therefore the enzyme luciferase enables, thereby enabling even more efficient protein splicing. Hence, for using the enzyme luciferase in the present invention, it is recommended that the enzyme be mutated into mutants R437C for Therefore, effective splicing may be induced by producing a mutant (R437C) in which the 437th arginine residue is substitute converted into cysteine, D436Y for a mutant (D436Y) in which the 436th asparagic acid residue is substituted converted into tyrosine, and I434A for a mutant (I434A) in which the 434th isoleucine residue is substituted converted into alanine to enable a more efficient protein splicing.

As mentioned hereinabove, the invention of this application provides the probe for protein-protein interaction analysis. Concretely Furthermore, one probe (e.g., probe a) is linked to the target one protein (protein A) of which the ability of whose interaction with another protein (protein B) is to be confirmed, while another the other probe (probe b "b") is linked to the other protein (protein B) whose interaction is to be confirmed, and the two probes are made to coexist in a system to a close proximity, thereby confirm the presence or absence enabling the analysis of interaction between proteins A and B by based on the above-mentioned principle and mechanism.

The proteins (4a, 4b) may be linked with the probes (1a, 1b) by any method, as long as the properties of the proteins and the probes are not affected. For example, any ordinary chemical, biochemical or genetic engineering strategies are applicable techniques can be employed.

The Any method of using the probe for protein-protein invention analysis of the invention may be employed for detecting and analyzing protein-protein interactions (e.g., through luminometry light emission) is not limited, and may be done by any ordinary means by using the probe for protein-protein interaction analysis of the invention. Experimental methods and detectors generally used in the field of chemistry and biochemistry may be employed. For example, a luminometer, which enable simple detection and analysis is preferable can be easily conducted by using a luminometer.

Embodiments of the invention are described in more detail in the following Examples with reference to the drawings attached hereto. It is needless to say, goes without saying that the invention is not limited to these Examples, and various changes and modifications may be made without limitation.

EXAMPLES

Example 1: Construction of Probe <1> for protein-protein interaction analysis

A yeast VMA1-derived intein (VDE) was used as As the intein site of probe <1> for protein-protein interaction analysis – yeast YMA10-derived intein (VDE) was used.

Of the 454 amino acids constituting VDE, Cys1 is taken as the first amino acid residue and Asn454 is taken as the last amino acid residue. For VDE, the The C-extein starts from Cys455, and the extein residue adjacent to Cys1 is numbered -1. The subsequent extein residues are numbered The numbers then increase as -2, -3, . . . toward the N-extein N Extein.

For the labeled site, an Enhanced a green fluorescent protein derived from Aequorea victoria light-emitting jelly fish (Aequorea victoria) (EGFP, described, for example, in *Current Biology* 1996, 6 (2), 178-182) was used.

In probe <1> for protein-protein interaction, as the

~~splicing site, the N-terminal half of EGFP requires one Cys residue, another a Gly residue upstream of the position of -1 position, and three hydrophobic amino acid residues at their positions -5, -4 and -3 positions have to be present in the N-terminal polypeptide of EGFP as the joining sections for splicing.~~

Accordingly, the amino acid residues between I124 and I129 ~~amino acid residues that are relatively stable in~~ where the N-terminal polypeptide structure of EGFP is relatively stable were substituted as follows (Fig. 2).

(1) Mutation of I129C and E125I mutation (hereinafter referred to as "m125") ~~in of EGFP showed fluorescence was conducted. The mutants showed fluorescence, and excitation and emission peaks equivalent to those of EGFP were observed at 488 nm and 510 nm were observed identical to those of EGFP.~~

(2) Fluorescence disappeared with L126Y mutation at m125 (hereinafter referred to as "m126"), although the expression level of the mutant did not vary was the same.

The results indicate indicated that L126Y mutation disables correct folding in the m129EGFP mutant was folded incorrectly and/or ligation of splicing sites of the fluorescent protein sites could not take place be linked due to L126Y mutation.

Accordingly, a m125EGFP mutant for which I129C and E125I mutation was performed was used as the labeled protein indicator protein site of the probe for protein-protein interaction. This labeled protein indicator protein site is hereinafter referred to as an EGFP mutant.

Example 2: Confirmation of splicing in single polypeptide

To confirm whether protein splicing occurred in a single polypeptide in which VDE was sandwiched between the N-terminal polypeptide and C-terminal havles polypeptide of the EGFP mutant, pGEX-NVC was expressed at 25°C in *E. coli*.

(1) Cells of *E. coli* DH5 α were ~~incubated~~ used to express a glutathion S-transferase fused protein(GST). A plasmid that covers the VDE region and the N- and C-terminal polypeptides of the EGFP mutant was fused to the GST gene under control or a tac promoter. ~~A cDNA that encodes the N-terminal polypeptide of m125 mutant GFP (1-128 amino acids), VDE (1-454), and C terminal polypeptide of m125 GFP (129-238 amino acids) was used.~~ This produced a chimera protein ~~consisting of comprising~~ GST (26 kDa), 125 residues (13 kDa) from the N-terminal polypeptide of the EGFP mutant ~~(13 kDa)~~, VDE (50 kDa) and the C-terminal polypeptide (14 kDa) of the mutant ~~(14 kDa)~~.

(2) The ~~resulting~~ protein obtained was extracted from the *E. coli* cells, purified, and identified by SDS-PAGE. The ~~labeled~~ protein indicator protein and 10 to 225-kDa marker (Novagen) were ~~electrophoresed added and electrophoresis was conducted on a~~ 12 to 15% SDS-PAGE gel. The gel was visualized by Coomassie Brilliant Blue staining.

~~For Western blotting[[,]] was conducted by using~~ anti-VDE polyclonal antibody, anti-His-labeled polyclonal antibody (Santa Crus Biotechnology), or anti-GFP monoclonal antibody (BioRad) ~~was used as the probe.~~

~~The All the~~ enzymes necessary for cloning were obtained from Takara Biomedical and used according to the manufacturer's instructions manual.

The PCR fragments were sequenced using a genetic analyzer, ABI310.

The main component of the crude product was a protein of about 50-kDa (Fig. 3b). ~~This result corresponds with It matched~~

the size of the extein linked to VDE (50-kDa) and GST. In other words, this component was ~~determined~~ assumed to be a fused protein of GST and the N-terminal polypeptide of the EGFP mutant (26-kDa + 13-kDa).

From this result, it was found that the 103-kDa precursor of the fused protein was split into the 50-kDa VDE and the 53-kDa GST-EGFP mutant fused protein.

In addition, the molecular ~~weight~~ weights of VDE and ~~that of the~~ GST-EGFP mutant fused protein were ~~measured through evaluated by~~ evaluated by Western blotting.

The anti-VDE and anti-GFP antibodies were specifically reacted with the excised 50-kDa intein (Fig. 3d) and the 53-kDa GST-EGFP mutant fused protein (Fig. 3e), respectively. ~~All The components observed in the vicinity of about 100-kDa of the unspliced precursor were analyzed using these antibodies.~~

The GST-EGFP mutant fused protein was further identified ~~by with a GST-affinity chromatography column. When the crude product was passed through an affinity column[[],]~~ and the proteins bound to the resin were ~~extracted taken out~~ using PreScission protease and then subjected to SDS-PAGE (Fig. 3c), a band was observed at 25-kDa. This result gave a 25-kDa band, which almost matches the same as the molecular weight of the EGFP mutant.

(3) The fluorescent spectrum of the protein purified with the affinity column was then measured. Next, the affinity column-purified protein was subjected to fluorescence spectrometry, for which The maximum wavelength of excitation and emission peaks appeared at was 488 nm and 510 nm, respectively, those results matching the values for EGFP. The data well corresponded to those of EGFP.

The above results indicated that the VDE located in the center of the single polypeptide was excised by splicing, ~~causing ligation of the N- and C-terminal polypeptides of the EGFP mutant~~

were linked by peptide bonding, resulting in the correct folding of and the obtained EGFP mutant folded correctly to form a fluorophore.

Example 3: Effect Effectiveness of probe <1> for protein-protein interaction analysis

(1) Cells of *E. coli* BL21(DE3)pLysS were incubated to produce a A recombinant fused protein labeled with His at its N-terminal was obtained by using cells of *E. coli* BL21(DE3).

In order to test protein-protein interaction, GST of pGEX-NVC pGEX_NVC was substituted with the His-label of a pET16-b vector, and its the splicing functional site was split to obtain pET-NVCASD(/) pET_NVCASD(/).

The splitting was realized by substituting the nonfunctional functionally unnecessary endonuclease motif that exists in the 185th to 389th amino acid region of the mutant with a cassette consisting of ((translation termination codon) - (Shine-Dalgarno sequence) - (translation initiation codon)).

The resulting plasmid pET-NVCASD(/) pET_NVCASD(/) was a two-gene operon essentially composed of gene 1 that encodes the N-terminal polypeptide of EGFP and VDE (pET-NVCASD(/) N_EGFP-VDE) and gene 2 that encodes the C-terminal polypeptide of VDE and EGFP (C-VDE-EGFP C_VDE-EGFP).

For In order to bring the N- and C-terminal polypeptides of VDE spatially close to each other in space during protein-protein interaction, a foldable bendable peptide linker having repetitive Gly-Asn sequences was introduced between the translation termination codon and the translation initiation codon (pET-NVCASD pET_NVCASD linker (/)).

In order to confirm the presence whether of specific protein-protein interaction that facilitates the splicing phenomenon to give providing EGFP in *E. coli* occurs, calmodulin (CaM) and its target peptide M13 were selected as model proteins.

The structure of CaM and that of M13 CaM and M13 are preferred, because the structures thereof have been clarified by NMR (Ikura, M., et al., *Science* 1992, 256, 632-638) [();] and, therefore, CaM and M13 are suitable because the distance between any amino acid in CaM and any amino acid in M13 is known.

(2) The recombinant plasmids pET-NVCASD(C/M) pET-NVCASD pET-NVCASD linker (C/M) and pET-NVCASD pET-NVCASD linker () were introduced into the cells of *E. coli* to obtain the corresponding fused protein.

For intracellular In order to induce the intracellular splicing, the protein expression was effected at 25°C for 12 hours at 25°C and the cells were thereafter stored at 4°C for 1 or 2 days at 4°C.

Gene 1 protein expression was confirmed by an anti-His labeled antibody (Fig. 5A). The main components analyzed by the antibody were proteins of 55-kDa (Fig. 5A a, d), and their size was about the same as that of the unspliced precursor of N-EGFP-VDE N-EGFP-VDE precursor (36-kDa) with CAM (17-kDa) or its linker (1-kDa).

As the control, When a lysate of *E. coli* having a pET-NVCASD pET-NVCASD linker () plasmid was subjected to SDS-PAGE as a control, which gave a single band of of N-EGFP-VDE N-EGFP-VDE (Fig. 5A, f) was obtained.

The expression level of the unspliced precursor proteins obtained from these three plasmids was almost the same.

On the other hand, the gene 2 protein product was identified by an anti-GFP monoclonal antibody, and was found to be the C-terminal polypeptide of EGFP.

In addition, in the cell that expressed pET-NVCASD(C/M) pET-NVCASD(C/M) and pET-NVCASD pET-NVCASD linker (C/M), the size of the expression protein obtained from the gene operon well corresponded agreed with that of the precursor protein of N-VDE-EGFP N-VDE-EGFP (20-kDa) with M13 (3-kDa) (Fig. 5B, a) or the

linker-containing M13 (4-kDa) (Fig. 5B, d). Similarly, the control plasmid expressed the expected protein (Fig. 5B, f). The expression levels of all three proteins obtained from operon II were the same.

Fig. 6 shows the ~~fluorescent~~ fluorescence spectra of lysates of *E. coli* cells, each carrying one of the above-mentioned plasmids. For the cells of *E. coli* comprising ~~pET-NVCASD~~ pET_NVCASD linker (/), no spectral change was observed. From the cells of *E. coli* ~~carrying~~ having ~~pET-NVCASD(C/M)~~ pET_NVCASD(C/M) and having co-expressed CaM and M13, a ~~certain~~ slight change in fluorescence emission at 510 nm was observed. On the other hand, ~~from the cells of E. Coli carrying N-EGFP-VDE in the co-expression of CaM and M13 where a bendable peptide linker was bonded to N-EGFP-VDE and C-VDE-EGFP with a peptide linker therebetween and having co-expressed CaM and M13, a significant change in the fluorescence emission was observed at 510 nm changed significantly.~~ The fluorescence intensity of the crude product from the cells of *E. coli* that had pET_NVCASD carries pET-NVCASD linker (C/M) was enough sufficient to differentiate those cells from the cells that carry the control plasmid, ~~pET-NVCASD~~ pET_NVCASD linker (/) or the plasmid not encoding the ~~foldable~~ bendable linker ~~pET-NVCASD~~ pET_NVCASD linker (C/M).

These results indicated that the CaM-M13 interaction resulted in trans splicing, ~~thereby inducing the and ligation of the two external regions of N- and C-terminal polypeptides of the EGFP mutant to form that formed~~ the EGFP fluorophore.

Example 4: Effect of linker in probe <1> for protein-protein interaction analysis

~~For protein To induce splicing, it is necessary that the N- and C-terminals of VDE be correctly folded. Correct Such correct~~ folding is attained only when the C-terminal ~~N-VDE~~ N_VDE is close

to the N-terminal C-VDE.

Regarding CaM and M13, the distance between the N-terminal of CaM and the C-terminal of M13 is 50 Å (from Brookhaven Protein Data Bank). This distance can be too large for N-VDE N VDE to be close to C-VDE, and the large distance was expected to interfere with and, therefore, was thought to inhibit the correct folding in *E. coli* carrying having the plasmid pET-NVCASD pET_NVCASD linker (C/M).

However, the plasmid pET-NVCASD actually increased a clear increase in the fluorescence intensity at 510 nm significantly. was observed in plasmid pET_NVCASD linker (C/M). This result confirms that the 10- and 9-amino acid linkers each that were linked to the N N and C-VDEs enabled the necessary foldability of VDE and realized good conformation for effective VDE splicing C-VDEs, respectively, provided bendability necessary for the VDE folding, and the conformation for effective VDE splicing was attained.

Example 5: Construction of probe <2> for protein-protein interaction analysis

This experiment was performed to demonstrate the construction of a plasmid in a host A plasmid was constructed as described below by using *E. coli* strain DH5 α as a host.

As the intein site for a probe <2>, DnaE derived from cyanobacterium *Synechocystis* sp. PCC6803 was chosen used as the intein site for a probe <2> for protein-protein interaction analysis. As the labeled protein indicator protein for the probe, a wild firefly-derived luciferase (Lum-F) (pLucAll) was split between the 437th amino acid and the 438th amino acid into and used as N-terminal (pLucN) and C-terminal (pLucC) segments.

Fig. 7 shows the structure of the luciferase and luciferase segments.

(1) Splitting of luciferase

First, the absence of enzymatic activity in pLucN and pLucC was confirmed by transitionally expressing the two luciferase segments in human insulin receptor-overexpressing Chinese hamster ovarian cells (CHO-HIR).

~~Concretely Thus, the CHO-HIR cells were transfected with 1 µg of each plasmid (pLucAll, pLucN, and pLucC) together with 0.01 µg of a control plasmid (pRL-TK), and incubated in a 12-well plate for 45 hours. After the incubation, the emission of the cells in each well was measured using with a luminometer.~~

As a ~~blank~~ control, the CHO-HIR cells containing no plasmid were incubated under the same ~~condition~~ conditions.

~~The In order to correct the transfection efficiency error in each well of the plate was corrected through, a dual-luciferase assay using a Renilla-derived luciferase (Lum-R) was conducted and the transfection efficiency in each well was corrected.~~

Fig. 8 shows the relative emission intensity (RLU) of each luciferase (or each luciferase segment).

The RLU of the CHO-HIR cells ~~with where~~ the wild firefly luciferase (LucAll) was expressed was 11.4 (Fig. 8a), ~~however, but~~ the RLU of the CHO-HIR cells where only LucN or LucC was expressed was 2.2×10^{-4} and 4.5×10^{-5} , respectively (Fig. 8b). ~~The Furthermore, the RLU (background) of the CHO-HIR cells transfected with Lum-R itself was 2.2×10^{-4} . This result indicates indicated that LucN or LucC alone shows showed no emission activity.~~

(2) Splitting of DnaE[[:]]

As mentioned above, *Synechocystis* sp. PCC6803-derived DnaE was split into ~~an~~ N-terminal segment of 123 amino acid residues and ~~a~~ C-terminal segment of 36 amino acid residues.

(3) Probe <2>[[:]] for protein-protein interaction

The above-described ~~segment~~ N-terminal of DnaE in (2) was

ligated with LucN in (1) while the C-terminal segment of DnaE in (2) was ligated with LucC in (1) [.] These segments, and then they were inserted into a multicloning site (MCS) of pIRES (Invitrogen), which is a bicistronic expression vector, pIRES (Invitrogen) at the multicloning site (MCS) to constructs obtain pIRES-DSL. The resulting plasmid (pIRES-DSL) had two MCSs: at the 3'-terminal of the N-terminal DnaE and at the 5'-terminal of the C-terminal DnaE (MCS-A and MCS-B, respectively), to which proteins. Proteins or protein domains which may interact with each other or which are to be tested for their interactivity are interaction can be introduced into those MCSs.

Fig. 7 shows the structure of pIRES-DSL.

Example 6: Effect Effectiveness 1 of probe <2> for protein-protein interaction analysis †

(1) This experiment was performed In order to confirm whether that probe <2> for protein-protein interaction analysis constructed in Example 5 is effective for luciferase acts effectively and induces splicing, using of luciferase, an oligopeptide (Y941) that contains comprises the 941st tyrosine residue of IRS-1 (which is known to participate physiologically in insulin signal transduction) and its target protein, SH2N domain derived from phosphatidylinositol 3-kinase-derived SH2N domain, which is a target protein thereof, were used (White, M. F., *Diabetologia* 1997, 40, S2-S17).

A plasmid pIRES-DSL (Y/S) with Y941 inserted into MCS-A and SH2N inserted into MCS-B[,] and a pRL-TK vector were transiently coexpressed in CHO-HIR cells.

Fig. 7 shows the structure of the pIRES-DSL (Y/S).

First, in a 6-well plate, CHO-HIR cells were transfected with 2 µg of the plasmid (pIRES-DSL (Y/S)) and 0.02 ng of the control plasmid pRL-TK in a 6-well plate. The cells were incubated for 45 hours, and then the culture was substituted with

an FBS-free solution of 1.0×10^{-7} M human insulin[[;]] and the CHO-HIR cells were ~~excited~~ stimulated through incubation (at 37°C) for 72 hours, 3 hours or 5 minutes. The insulin solution of the cells incubated for 5 minutes ~~were~~ was further then incubated for 175 minutes in an FCS-free culture ~~with~~ containing no insulin ~~therein~~.

Fig. 9 shows the ~~relative~~ emission intensity, which is represented by RLU, of the ~~insulin-excited~~ cells after the insulin stimulation.

The corrected luciferase activity of the CHO-HIR cells ~~excited~~ stimulated with insulin for 72 hours, 3 hours and 5 minutes ~~were~~ was 1.0, 0.73, and 0.18, respectively. The ~~relative~~ emission intensity (against background) of the cells incubated in the insulin-free culture was 0.15, and ~~this is nearly~~ was found to be almost the same as that of the cells ~~excited~~ stimulated with insulin for 5 minutes.

~~In addition, it is understood that the relative~~
Furthermore, the emission intensity, which was represented by RLU, of the cells ~~excited~~ stimulated with insulin for 3 hours or longer ~~is~~ was found to be more than 4 times that of the background value. The kinase activity of the cells ~~tested~~ observed herein results resulted from the specific interaction between SH2N and the peptide Y941 (the 941 941st amino acid of which is tyrosine) phosphorylated with an insulin receptor, and SH2N. This ~~enabled~~ the efficient in-trans folding and splicing of DnaE occurred effectively in probe <2>[[,]] for protein-protein interaction, LucN and LucC to were linked, and luciferase emission reproduction was reproduced.

~~Next, an~~ An Y941 mutant ~~(wherein, in which~~ the tyrosine residue in Y941 was substituted with an alanine residue that ~~is~~ was not phosphorylated with the insulin receptor[[,]] was then constructed and tested in the same manner as described above.

The results are shown in Fig. 10.

The In the CHO-HIR cells where the Y941 mutant was expressed, the corrected luciferase emission of the Y941 mutant-
expressing CHO-HIR cells was on the same level as that of the background. This confirms confirmed that the tyrosine phosphorylation in Y941 led to the protein- protein interaction between Y941 and SH2N with SH2N was due to phosphorylation of tyrosine in Y941.

Example 7: Effect Effectiveness 2 of probe <2> 2 for protein- protein interaction analysis

Probe <2> constructed in Example 5 was tested for quantitative Quantitative analysis of insulin-induced protein- protein interaction in CHO-HIR cells was conducted by using probe <2> for protein-protein interaction analysis constructed in Example 5.

First the insulin concentration dependence of the RLU value was confirmed by the following procedure: pIRES-DSL (Y/S) and pRL-TK were coexpressed in After the CHO-HIR cells were coexpressed by pIRES-DSL (Y/S) and pRL-TK in the same manner as described above, and the cells were then excited stimulated at 37°C for 3 hours with insulin of various concentrations between 1.0 x 10⁻¹³ and 1.0 x 10⁻⁷ M, at 37°C for 3 hours by the above-described method.

Fig. 11 shows the emission intensity in RLU.

The results indicate that the The emission intensity of the cells increases was confirmed to increase with insulin concentration.

INDUSTRIAL APPLICABILITY

As described in detail hereinabove, the present invention provides a probe for protein-protein interaction analysis that enables accurate and simple analysis of protein-protein interactions. The invention also provides a method for the

analysis of protein-protein interactions using the probe for protein-protein interaction analysis.

The present method for analysis of protein-protein interactions does not require any reporter gene or substrate as in the conventional methods, and enables simple and accurate analysis ~~or protein-protein interactions.~~ The method and the probe are applicable even to mammal cells for high-sensitivity analysis of protein-protein interactions in the cells. Furthermore, protein-protein interactions can be also analyzed with high-sensitivity in mammal cells.

Accordingly, the probe for protein-protein interaction analysis ~~of~~ in accordance with the present invention enables fast and simple analysis of the mechanisms of protein-protein interactions in various organisms, such as protein-protein interactions in cell membranes and receptor activation by hormones ~~in~~ inside the cells, ~~in a simplified manner that requires little time.~~

ABSTRACT

A probe for protein-protein interaction analysis ~~applicable to suitable for analyzing protein-protein interactions of various proteins, which enables simple analysis in with high accuracy and in a simple manner~~ [1,2] and a method for analyzing interaction of two proteins ~~is provided by using the probe. Protein With the probe, protein splicing from the probes is~~ caused by protein-protein interaction, and a physicochemically or biochemically detectable protein is regenerated.